# Genome-wide association study of germline variants and breast cancer-specific mortality

Maria Escala-Garcia[1]*, Qi Guo[2]*, Thilo Dörk[3], Sander Canisius[1, 4], Renske Keeman[1], Joe Dennis[5], Jonathan Beesley[6], Julie Lecarpentier[5], Manjeet K. Bolla[5], Qin Wang[5], Jean Abraham[7-9], Irene L. Andrulis[10, 11], Hoda Anton-Culver[12], Volker Arndt[13], Paul L. Auer[14, 15], Matthias W. Beckmann[16], Sabine Behrens[17], Javier Benitez[18, 19], Marina Bermisheva[20], Leslie Bernstein[21], Carl Blomqvist[22, 23], Bram Boeckx[24, 25], Stig E. Bojesen[26-28], Bernardo Bonanni[29], Anne-Lise Børresen-Dale[30, 31], Hiltrud Brauch[32-34], Hermann Brenner[13, 34, 35], Adam Brentnall[36], Louise Brinton[37], Per Broberg[38], Ian W. Brock[39], Sara Y. Brucker[40], Barbara Burwinkel[41, 42], Carlos Caldas[8, 9, 43], Trinidad Caldés[44], Daniele Campa[17, 45], Federico Canzian[42], Angel Carracedo[46-48], Brian D. Carter[49], Jose E. Castelao[50], Jenny Chang-Claude[17, 51], Stephen J. Chanock[37], Georgia Chenevix-Trench[6], Ting-Yuan David Cheng[52], Suet-Feung Chin[53], Christine L. Clarke[54], NBCS Collaborators[30-126], Emilie Cordina-Duverger[55], Fergus J. Couch[56], David G. Cox[57, 58], Angela Cox[39], Simon S. Cross[59], Kamila Czene[60], Mary B. Daly[61], Peter Devilee[62, 63], Janet A. Dunn[64], Alison M. Dunning[7], Lorraine Durcan[65, 66], Miriam Dwek[67], Helena M. Earl[9, 68], Arif B. Ekici[69], A. Heather Eliassen [70, 71], Carolina Ellberg[38], Christoph Engel[72, 73], Mikael Eriksson[60], D. Gareth Evans[74, 75], Jonine Figueroa[37, 77, 78], Dieter Flesch-Janys[79, 80], Henrik Flyger[81], Marike Gabrielson[60], Manuela Gago-Dominguez[46, 82], Eva Galle[24, 25], Susan M. Gapstur[49], Montserrat García-Closas[37, 83], José A. García-Sáenz[44], Mia M. Gaudet[49], Angela George[84, 85], Vassilios Georgoulias[86], Graham G. Giles[87-89], Gord Glendon[10], David E. Goldgar[90], Anna González-Neira[18], Grethe I. Grenaker Alnæs[30], Mervi Grip[91], Pascal Guénel[55], Lothar Haeberle[92], Eric Hahnen[93, 94], Christopher A. Haiman[95], Niclas Håkansson[96], Per Hall[60, 97], Ute Hamann[98], Susan Hankinson[99], Elaine F. Harkness[100-102], Patricia A. Harrington[7], Steven N. Hart[103], Jaana M. Hartikainen[104-106], Alexander Hein[16], Peter Hillemanns[3], Louise Hiller[64], Bernd Holleczek[107], Antoinette Hollestelle[108], Maartje J.

26Hooning[108], Robert N. Hoover[37], John L. Hopper[88], Anthony Howell[109], Guanmengqian
27Huang[98], Keith Humphreys[60], David J. Hunter[71, 110, 111], Wolfgang Janni[112], Esther M. John[113-
28115], Michael E. Jones[83], Arja Jukkola-Vuorinen[116], Audrey Jung[17], Rudolf Kaaks[17], Maria
29Kabisch[98], Katarzyna Kaczmarek[117], Michael J. Kerin[118], Sofia Khan[119], Elza Khusnutdinova[20,
30120], Johanna I. Kiiski[119], Cari M. Kitahara[121], Julia A. Knight[122, 123], Yon-Dschun Ko[124], Linetta B.
31Koppert[125], Veli-Matti Kosma[104-106], Peter Kraft[71, 110], Vessela N. Kristensen[30, 31], Ute Krüger[38],
32Tabea Kühl[51], Diether Lambrechts[24, 25], Loic Le Marchand[126], Eunjung Lee[95], Flavio
33Lejbkowicz[127], Lian Li[128], Annika Lindblom[129], Sara Lindström[130, 131], Martha Linet[121], Jolanta
34Lissowska[132], Wing-Yee Lo[32, 33], Sibylle Loibl[133], Jan Lubiński[117], Michael P. Lux[92], Robert J.
35MacInnis [87, 88], Melanie Maierthaler[42], Tom Maishman[65, 66], Enes Makalic[88], Arto
36Mannermaa[104-106], Mehdi Manoochehri[98], Siranoush Manoukian[134], Sara Margolin[135], Maria
37Elena Martinez[82, 136], Dimitrios Mavroudis[86], Catriona McLean[137], Alfons Meindl[138], Pooja
38Middha[17, 139], Nicola Miller[118], Roger L. Milne[87, 88], Fernando Moreno[44], Anna Marie Mulligan[140,
39141], Claire Mulot[142], Rami Nassir[143], Susan L. Neuhausen[21], William T. Newman[74, 75], Sune F.
40Nielsen[26, 27], Børge G. Nordestgaard[26-28], Aaron Norman[103], Håkan Olsson[38], Nick Orr[144], V.
41Shane Pankratz[145], Tjoung-Won Park-Simon[3], Jose I.A. Perez[146], Clara Pérez-Barrios[147],
42Paolo Peterlongo[148], Christos Petridis[149], Paul D.P. Pharoah[5, 7], Mila Pinchev[127], Karoliona
43Prajzendanc[117], Ross Prentice[14], Nadege Presneau[67], Darya Prokofieva[120], Katri Pylkäs[150, 151],
44Brigitte Rack[138], Paolo Radice[152], Dhanya Ramachandran[3], Gadi Rennert[127], Hedy S.
45Rennert[127], Valerie Rhenius[7], Atocha Romero[147], Rebecca Roylance[153], Emmanouil
46Saloustros[154], Elinor J. Sawyer[149], Daniel F. Schmidt[88], Rita K. Schmutzler[93, 94], Andreas
47Schneeweiss[41, 156], Minouk J. Schoemaker[84], Fredrick Schumacher[157], Lukas Schwentner[112],
48Rodney J. Scott[158-161], Christopher Scott[103], Caroline Seynaeve[108], Mitul Shah[7], Jacques
49Simard[162], Ann Smeets[163], Christof Sohn[156], Melissa C. Southey[164, 165], Anthony J. Swerdlow[84,
50166], Aline Talhouk[167-169], Rulla M. Tamimi[70, 71, 110], William J. Tapper[170], Manuel R. Teixeira[171,
51172], Maria Tengström[104, 173, 174], Mary Beth Terry[175], Kathrin Thöne[51], Rob A.E.M. Tollenaar[176],

Ian Tomlinson[177, 178], Diana Torres[98, 179], Thérèse Truong[55], Constance Turman[71], Clare Turnbull[84], Hans-Ulrich Ulmer[180], Michael Untch[181], Celine Vachon[103], Christi J. van Asperen[182], Ans M.W. van den Ouweland[183], Elke M. van Veen[74, 75], Camilla Wendt[184], Alice S. Whittemore[114, 115], Walter Willett [71, 185, 186], Robert Winqvist[150, 151], Alicja Wolk[187], Xiaohong R. Yang[37], Yan Zhang[13, 34], Douglas F. Easton[5, 7], Peter A. Fasching[16, 76], Heli Nevanlinna[119], Diana M. Eccles[66], Paul D.P. Pharoah[5, 7] and Marjanka K. Schmidt[1, 155].

*Shared first authorship

**Correspondence to:** Qi Guo

[1] The Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Division of Molecular Pathology, Amsterdam, The Netherlands.
[2] University of Cambridge, Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, Cambridge, UK.
[3] Hannover Medical School, Gynaecology Research Unit, Hannover, Germany.
[4] The Netherlands Cancer Institute - Antoni van Leeuwenhoek hospital, Division of Molecular Carcinogenesis, Amsterdam, The Netherlands.
[5] University of Cambridge, Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, Cambridge, UK.
[6] QIMR Berghofer Medical Research Institute, Department of Genetics and Computational Biology, Brisbane, Queensland, Australia.
[7] University of Cambridge, Centre for Cancer Genetic Epidemiology, Department of Oncology, Cambridge, UK.
[8] Cambridge Experimental Cancer Medicine Centre, Cambridge, UK.
[9] University of Cambridge NHS Foundation Hospitals, Cambridge Breast Unit and NIHR Cambridge Biomedical Research Centre, Cambridge, UK.
[10] Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Fred A. Litwin Center for Cancer Genetics, Toronto, ON, Canada.
[11] University of Toronto, Department of Molecular Genetics, Toronto, ON, Canada.
[12] University of California Irvine, Department of Epidemiology, Genetic Epidemiology Research Institute, Irvine, CA, USA.
[13] German Cancer Research Center (DKFZ), Division of Clinical Epidemiology and Aging Research, Heidelberg, Germany.
[14] Fred Hutchinson Cancer Research Center, Cancer Prevention Program, Seattle, WA, USA.
[15] University of Wisconsin-Milwaukee, Zilber School of Public Health, Milwaukee, WI, USA.
[16] University Hospital Erlangen, Friedrich-Alexander-University Erlangen-Nuremberg, Department of Gynecology and Obstetrics, Comprehensive Cancer Center ER-EMN, Erlangen, Germany.


3

[17] German Cancer Research Center (DKFZ), Division of Cancer Epidemiology, Heidelberg, Germany.
[18] Spanish National Cancer Research Centre (CNIO), Human Cancer Genetics Programme, Madrid, Spain.
[19] Biomedical Network on Rare Diseases (CIBERER), Madrid, Spain.
[20] Ufa Scientific Center of Russian Academy of Sciences, Institute of Biochemistry and Genetics, Ufa, Russia.
[21] Beckman Research Institute of City of Hope, Department of Population Sciences, Duarte, CA, USA.
[22] University of Helsinki, Department of Oncology, Helsinki University Hospital, Helsinki, Finland.
[23] Örebro University Hospital, Department of Oncology, Örebro, Sweden.
[24] VIB, VIB Center for Cancer Biology, Leuven, Belgium.
[25] University of Leuven, Laboratory for Translational Genetics, Department of Human Genetics, Leuven, Belgium.
[26] Copenhagen University Hospital, Copenhagen General Population Study, Herlevand Gentofte Hospital, Herlev, Denmark.
[27] Copenhagen University Hospital, Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Herlev, Denmark.
[28] University of Copenhagen, Faculty of Health and Medical Sciences, Copenhagen, Denmark.
[29] Istituto Europeo di Oncologia, Division of Cancer Prevention and Genetics, Milan, Italy.
[30] Oslo University Hospital-Radiumhospitalet, Department of Cancer Genetics, Institute for Cancer Research, Oslo, Norway.
[31] University of Oslo, Institute of Clinical Medicine, Faculty of Medicine, Oslo, Norway.
[32] Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, Germany.
[33] University of Tübingen, Tübingen, Germany.
[34] German Cancer Research Center (DKFZ), German Cancer Consortium (DKTK), Heidelberg, Germany.
[35] German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Division of Preventive Oncology, Heidelberg, Germany.
[36] Queen Mary University of London, Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, London, UK.
[37] National Cancer Institute, Division of Cancer Epidemiology and Genetics, Bethesda, MD, USA.
[38] Lund University, Department of Cancer Epidemiology, Clinical Sciences, Lund, Sweden.
[39] University of Sheffield, Sheffield Institute for Nucleic Acids (SInFoNiA), Department of Oncology and Metabolism, Sheffield, UK.
[40] University of Tübingen, Department of Gynecology and Obstetrics, Tübingen, Germany.
[41] University of Heidelberg, Department of Obstetrics and Gynecology, Heidelberg, Germany.
[42] German Cancer Research Center (DKFZ), Molecular Epidemiology Group, C080, Heidelberg, Germany.


4

[43] The Institute of Cancer Research, Section of Cancer Genetics, London, UK.
[44] Instituto de Investigaci—n Sanitaria San Carlos (IdISSC), Centro Investigaci—n BiomŽdica en Red de C‡ncer (CIBERONC), Medical Oncology Department, Hospital Cl'nico San Carlos, Madrid, Spain.
[45] University of Pisa, Department of Biology, Pisa, Italy.
[46] Instituto de Investigaci—n Sanitaria de Santiago de Compostela (IDIS), Genomic Medicine Group, Galician Foundation of Genomic Medicine, SERGAS, Santiago de Compostela, Spain.
[47] Universidad de Santiago de Compostela, Centro de Investigaci—n en Red de Enfermedades Raras (CIBERER), Santiago De Compostela, Spain.
[48] King Abdulaziz University, Center of Excellence in Genomic Medicine, Jeddah, KSA.
[49] American Cancer Society, Epidemiology Research Program, Atlanta, GA, USA.
[50] Instituto de Investigacion Sanitaria Galicia Sur (IISGS), Xerencia de Xestion Integrada de Vigo-SERGAS, Oncology and Genetics Unit, Vigo, Spain.
[51] University Medical Center Hamburg-Eppendorf, Cancer Epidemiology Group, University Cancer Center Hamburg (UCCH), Hamburg, Germany.
[52] Roswell Park Cancer Institute, Division of Cancer Prevention and Control, Buffalo, NY, USA.
[53] University of Cambridge, Cancer Research UK Cambridge Institute, Cambridge, UK.
[54] University of Sydney, Westmead Institute for Medical Research, Sydney, New South Wales, Australia.
[55] INSERM, University Paris-Sud, University Paris-Saclay, Cancer & Environment Group, Center for Research in Epidemiology and Population Health (CESP), Villejuif, France.
[56] Mayo Clinic, Department of Laboratory Medicine and Pathology, Rochester, MN, USA.
[57] Imperial College London, Department of Epidemiology and Biostatistics, School of Public Health, London, UK.
[58] Cancer Research Center of Lyon, INSERM U1052, Lyon, France.
[59] University of Sheffield, Academic Unit of Pathology, Department of Neuroscience, Sheffield, UK.
[60] Karolinska Institutet, Department of Medical Epidemiology and Biostatistics, Stockholm, Sweden.
[61] Fox Chase Cancer Center, Department of Clinical Genetics, Philadelphia, PA, USA.
[62] Leiden University Medical Center, Department of Pathology, Leiden, The Netherlands.
[63] Leiden University Medical Center, Department of Human Genetics, Leiden, The Netherlands.
[64] University of Warwick, Warwick Clinical Trials Unit, Coventry, UK.
[65] University of Southampton, Southampton Clinical Trials Unit, Faculty of Medicine, Southampton, UK.
[66] University of Southampton, Cancer Sciences Academic Unit, Faculty of Medicine, Southampton, UK.

[67] University of Westminster, Department of Biomedical Sciences, Faculty of Science and Technology, London, UK.
[68] University of Cambridge, Department of Oncology, Cambridge, UK.
[69] Friedrich-Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, Institute of Human Genetics, University Hospital Erlangen, Erlangen, Germany.
[70] Harvard Medical School, Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA.
[71] Harvard T.H. Chan School of Public Health, Department of Epidemiology, Boston, MA, USA.
[72] University of Leipzig, Institute for Medical Informatics, Statistics and Epidemiology, Leipzig, Germany.
[73] University of Leipzig, LIFE - Leipzig Research Centre for Civilization Diseases, Leipzig, Germany.
[74] University of Manchester, Manchester Academic Health Science Centre, Division of Evolution and Genomic Medicine, School of Biological Sciences, Faculty of Biology, Medicine and Health, Manchester, UK.
[75] St Marys Hospital, Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester Centre for Genomic Medicine, Manchester, UK.
[76] University of California at Los Angeles, David Geffen School of Medicine, Department of Medicine Division of Hematology and Oncology, Los Angeles, CA, USA.
[77] The University of Edinburgh Medical School, Usher Institute of Population Health Sciences and Informatics, Edinburgh, UK.
[78] Cancer Research UK Edinburgh Centre, Edinburgh, UK.
[79] University Medical Center Hamburg-Eppendorf, Institute for Medical Biometrics and Epidemiology, Hamburg, Germany.
[80] University Medical Center Hamburg-Eppendorf, Department of Cancer Epidemiology, Clinical Cancer Registry, Hamburg, Germany.
[81] Copenhagen University Hospital, Department of Breast Surgery, Herlev and Gentofte Hospital, Herlev, Denmark.
[82] University of California San Diego, Moores Cancer Center, La Jolla, CA, USA.
[83] Institute of Cancer Research, Division of Genetics and Epidemiology, London, UK.
[84] The Institute of Cancer Research, Division of Genetics and Epidemiology, London, UK.
[85] The Royal Marsden NHS Foundation Trust, Cancer Genetics Unit, London, UK.
[86] University Hospital of Heraklion, Department of Medical Oncology, Heraklion, Greece.
[87] Cancer Council Victoria, Cancer Epidemiology & Intelligence Division, Melbourne, Victoria, Australia.
[88] The University of Melbourne, Melbourne School of Population and Global Healthm, Centre for Epidemiology and Biostatistics, Melbourne, Victoria, Australia.
[89] Monash University, Department of Epidemiology and Preventive Medicine, Melbourne, Victoria, Australia.


6

[90] Huntsman Cancer Institute, University of Utah School of Medicine, Department of Dermatology, Salt Lake City, UT, USA.
[91] University of Oulu, Department of Surgery, Oulu University Hospital, Oulu, Finland.
[92] Friedrich-Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, Department of Gynaecology and Obstetrics, University Hospital Erlangen, Erlangen, Germany.
[93] University Hospital of Cologne, Center for Hereditary Breast and Ovarian Cancer, Cologne, Germany.
[94] University of Cologne, Center for Molecular Medicine Cologne (CMMC), Cologne, Germany.
[95] University of Southern California, Department of Preventive Medicine, Keck School of Medicine, Los Angeles, CA, USA.
[96] Karolinska Institutet, Institute of Environmental Medicine, Stockholm, Sweden.
[97] South General Hospital, Department of Oncology, Stockholm, Sweden.
[98] German Cancer Research Center (DKFZ), Molecular Genetics of Breast Cancer, Heidelberg, Germany.
[99] University of Massachusetts, Amherst, Department of Biostatistics & Epidemiology, Amherst, MA, USA.
[100] University of Manchester, Manchester Academic Health Science Centre, Division of Informatics, Imaging and Data Sciences, Faculty of Biology, Medicine and Health, Manchester, UK.
[101] Wythenshawe Hospital, Manchester University NHS Foundation Trust, Nightingale Breast Screening Centre, Manchester, UK.
[102] Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, NIHR Manchester Biomedical Research Unit, Manchester, UK.
[103] Mayo Clinic, Department of Health Sciences Research, Rochester, MN, USA.
[104] University of Eastern Finland, Translational Cancer Research Area, Kuopio, Finland.
[105] University of Eastern Finland, Institute of Clinical Medicine, Pathology and Forensic Medicine, Kuopio, Finland.
[106] Kuopio University Hospital, Imaging Center, Department of Clinical Pathology, Kuopio, Finland.
[107] Saarland Cancer Registry, Saarbruecken, Germany.
[108] Erasmus MC Cancer Institute, Department of Medical Oncology, Family Cancer Clinic, Rotterdam, The Netherlands.
[109] University of Manchester, Institute of Cancer studies, Manchester, UK.
[110] Harvard T.H. Chan School of Public Health, Program in Genetic Epidemiology and Statistical Genetics, Boston, MA, USA.
[111] University of Oxford, Nuffield Department of Population Health, Oxford, UK.
[112] Cancer Research UK Edinburgh Centre
[113] Cancer Prevention Institute of California, Department of Epidemiology, Fremont, CA, USA.
[114] Stanford University School of Medicine, Department of Health Research and Policy - Epidemiology, Stanford, CA, USA.


7

[115] Stanford University School of Medicine, Department of Biomedical Data Science, Stanford, CA, USA.
[116] University of Oulu, Department of Oncology, Oulu University Hospital, Oulu, Finland.
[117] Pomeranian Medical University, Department of Genetics and Pathology, Szczecin, Poland.
[118] National University of Ireland, Surgery, School of Medicine, Galway, Ireland.
[119] University of Helsinki, Department of Obstetrics and Gynecology, Helsinki University Hospital, Helsinki, Finland.
[120] Bashkir State University, Department of Genetics and Fundamental Medicine, Ufa, Russia.
[121] National Cancer Institute, Radiation Epidemiology Branch, Division of Cancer Epidemiology and Genetics, Bethesda, MD, USA.
[122] Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Prosserman Centre for Population Health Research, Toronto, ON, Canada.
[123] University of Toronto, Division of Epidemiology, Dalla Lana School of Public Health, Toronto, ON, Canada.
[124] Johanniter Krankenhaus, Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Bonn, Germany.
[125] Erasmus MC Cancer Institute, Department of Surgical Oncology, Family Cancer Clinic, Rotterdam, The Netherlands.
[126] University of Hawaii Cancer Center, Epidemiology Program, Honolulu, HI, USA.
[127] Carmel Medical Center and Technion Faculty of Medicine, Clalit National Cancer Control Center, Haifa, Israel.
[128] Tianjin Medical University Cancer Institute and Hospital, Department of Epidemiology, Tianjin, China.
[129] Karolinska Institutet, Department of Molecular Medicine and Surgery, Stockholm, Sweden.
[130] University of Washington School of Public Health, Department of Epidemiology, Seattle, WA, USA.
[131] Fred Hutchinson Cancer Research Center, Public Health Sciences Division, Seattle, WA, USA.
[132] M. Sklodowska-Curie Cancer Center- Oncology Institute, Department of Cancer Epidemiology and Prevention, Warsaw, Poland.
[133] GmbH, German Breast Group, Neu Isenburg, Germany.
[134] Fondazione IRCCS (Istituto Di Ricovero e Cura a Carattere Scientifico) Istituto Nazionale dei Tumori (INT), Unit of Medical Genetics, Department of Medical Oncology and Hematology, Milan, Italy.
[135] Karolinska Institutet, Department of Clinical Science and Education, Sšdersjukhuset, Stockholm, Sweden.
[136] University of California San Diego, Department of Family Medicine and Public Health, La Jolla, CA, USA.
[137] The Alfred Hospital, Anatomical Pathology, Melbourne, Victoria, Australia.
[138] Ludwig Maximilian University of Munich, Department of Gynecology and Obstetrics, Munich, Germany.
[139] University of Heidelberg, Faculty of Medicine, Heidelberg, Germany.


8

[140] University of Toronto, Department of Laboratory Medicine and Pathobiology, Toronto, ON, Canada.
[141] University Health Network, Laboratory Medicine Program, Toronto, ON, Canada.
[142] INSERM UMR-S1147, Université Paris Sorbonne Cité, Paris, France.
[143] University of California Davis, Department of Biochemistry and Molecular Medicine, Davis, CA, USA.
[144] Queen's University Belfast, Centre for Cancer Research and Cell Biology, Belfast, Ireland, UK.
[145] University of New Mexico, University of New Mexico Health Sciences Center, Albuquerque, NM, USA.
[146] Hospital Monte Naranco, Servicio de Cirug’a General y Especialidades, Oviedo, Spain.
[147] Hospital Universitario Puerta de Hierro, Medical Oncology Department, Madrid, Spain.
[148] The FIRC (Italian Foundation for Cancer Research) Institute of Molecular Oncology, IFOM, Milan, Italy.
[149] King's College London, Research Oncology, Guy's Hospital, London, UK.
[150] University of Oulu, Laboratory of Cancer Genetics and Tumor Biology, Cancer and Translational Medicine Research Unit, Biocenter Oulu, Oulu, Finland.
[151] Northern Finland Laboratory Centre Oulu, Laboratory of Cancer Genetics and Tumor Biology, Oulu, Finland.
[152] Fondazione IRCCS (Istituto Di Ricovero e Cura a Carattere Scientifico) Istituto Nazionale dei Tumori (INT), Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Research, Milan, Italy.
[153] UCLH Foundation Trust, Department of Oncology, London, UK.
[154] University Hospital of Larissa, Department of Oncology, Larissa, Greece.
[155] The Netherlands Cancer Institute - Antoni van Leeuwenhoek hospital, Division of Psychosocial Research and Epidemiology, Amsterdam, The Netherlands.
[156] University of Heidelberg, National Center for Tumor Diseases, Heidelberg, Germany.
[157] Case Western Reserve University, Department of Population and Quantitative Health Sciences, Cleveland, OH, USA.
[158] John Hunter Hospital, Division of Molecular Medicine, Pathology North, Newcastle, New South Wales, Australia.
[159] University of Newcastle, Discipline of Medical Genetics, School of Biomedical Sciences and Pharmacy, Faculty of Health, Callaghan, New South Wales, Australia.
[160] John Hunter Hospital, Hunter Medical Research Institute, Newcastle, New South Wales, Australia.
[161] University of Newcastle, Centre for Information Based Medicine, Callaghan, New South Wales, Australia.
[162] Centre Hospitalier Universitaire de Québec - Université Laval Research Center, Genomics Center, Québec City, QC, Canada.
[163] University Hospitals Leuven, Department of Surgical Oncology, Leuven, Belgium.


9

[164] Monash University, Precision Medicine, School of Clinical Sciences at Monash Health, Clayton, Victoria, Australia.
[165] The University of Melbourne, Department of Clinical Pathology, Melbourne, Victoria, Australia.
[166] The Institute of Cancer Research, Division of Breast Cancer Research, London, UK.
[167] BC Cancer Agency and University of British Columbia, British Columbia's Ovarian Cancer Research (OVCARE) Program, Vancouver General Hospital, Vancouver, BC, Canada.
[168] University of British Columbia, Department of Pathology and Laboratory Medicine, Vancouver, BC, Canada.
[169] University of British Columbia, Department of Obstetrics and Gynaecology, Vancouver, BC, Canada.
[170] University of Southampton, Faculty of Medicine, Southampton, UK.
[171] Portuguese Oncology Institute, Department of Genetics, Porto, Portugal.
[172] University of Porto, Biomedical Sciences Institute (ICBAS), Porto, Portugal.
[173] Kuopio University Hospital, Cancer Center, Kuopio, Finland.
[174] University of Eastern Finland, Institute of Clinical Medicine, Oncology, Kuopio, Finland.
[175] Columbia University, Department of Epidemiology, Mailman School of Public Health, New York, NY, USA.
[176] Leiden University Medical Center, Department of Surgery, Leiden, The Netherlands.
[177] University of Birmingham, Institute of Cancer and Genomic Sciences, Birmingham, UK.
[178] University of Oxford, Wellcome Trust Centre for Human Genetics and Oxford NIHR Biomedical Research Centre, Oxford, UK.
[179] Pontificia Universidad Javeriana, Institute of Human Genetics, Bogota, Colombia.
[180] Frauenklinik der Stadtklinik Baden-Baden, Baden-Baden, Germany.
[181] Helios Clinics Berlin-Buch, Department of Gynecology and Obstetrics, Berlin, Germany.
[182] Leiden University Medical Center, Department of Clinical Genetics, Leiden, The Netherlands.
[183] Erasmus University Medical Center, Department of Clinical Genetics, Rotterdam, The Netherlands.
[184]
[185] Harvard T.H. Chan School of Public Health, Department of Nutrition, Boston, MA, USA.
[186] Brigham and Women's Hospital and Harvard Medical School, Channing Division of Network Medicine, Boston, MA, USA.
[187] Karolinska Institutet, Department of Environmental Medicine, Division of Nutritional Epidemiology, Stockholm, Sweden.

## Abstract


10

**Background:**

We examined the associations between germline variants and breast cancer mortality using a large meta-analysis of women of European ancestry.

**Methods:**

Meta-analyses included summary estimates based on Cox models of twelve datasets using ~10.4 million variants for 96,661 women with breast cancer and 7,697 events (breast cancer-specific deaths). Estrogen receptor (ER)-specific analyses were based on 64,171 ER-positive (4,116) and 16,172 ER-negative (2,125) patients. We evaluated the probability of a signal to be a true positive using the Bayesian false discovery probability (BFDP).

**Results:**

We did not find any variant associated with breast cancer-specific mortality at $P<5\times10^{-8}$. For ER-positive disease, the most significantly associated variant was chr7:rs4717568 (BFDP=7%, $P=1.28\times10^{-7}$, hazard ratio [HR]=0.88, 95% confidence interval [CI]=0.84-0.92); the closest gene is *AUTS2*. For ER-negative disease, the most significant variant was chr7:rs67918676 (BFDP=11%, $P=1.38\times10^{-7}$, HR=1.27, 95% CI=1.16-1.39); located within a long intergenic non-coding RNA gene (AC004009.3), close to the *HOXA* gene cluster.

**Conclusions:**

We uncovered germline variants on chromosome 7 at BFDP<15% close to genes for which there is biological evidence related to breast cancer outcome. However, the paucity of variants associated with mortality at genome-wide significance underpins the challenge in providing genetic-based individualised prognostic information for breast cancer patients.

## Main

Breast cancer is the most common cancer in the Western world and accounts for 15% of cancer-related deaths in women, with about 522,000 deaths worldwide in 2012[1]. Survival


11

after a diagnosis of breast cancer varies considerably between patients even with closely matching tumour characteristics. Models that predict the likelihood of survival after breast cancer treatment use tumour and treatment data, but currently do not take host factors into account. The identification of prognostic and predictive biomarkers inherent in the germline of the patients rather than the tumour could pinpoint mechanisms of tumour progression and help with treatment stratification to increase therapeutic benefit. Such markers include inherited genetic variation, as there is evidence for heritability of breast cancer-specific mortality in affected first-degree relatives[2–5]. Germline variation may affect prognosis by affecting tumour biology, since such variants are known to be associated with risk of specific breast tumour subtypes, particularly those defined by hormone receptor status, and have different outcomes[6–8]. Germline genotype could also affect the efficacy of adjuvant drug therapies[9,10] or might condition the host tumour environment via vascularisation[11,12], metastatic pattern[13,14], stroma-tumour interaction[15,16] and immune surveillance[17,18].

The association between common germline genetic variation and breast cancer-specific mortality has been examined in many candidate gene studies[5,9,26–35,14,36,19–25], as well as in moderate-sized genome-wide association studies (GWAS)[37–41]. However, it has been difficult link GWAS results to plausible candidate genes and few have been convincingly replicated[36,42]. Large studies with long follow-up and reliable data on known prognostic factors are required if novel alleles associated with prognosis in breast cancer are to be identified at a level of genome-wide significance. In the present work, we pooled genotype data from multiple breast cancer GWAS discovery and replication efforts[43,44] with new genotype data obtained from a large breast cancer series genotyped using the OncoArray chip[45,46]. We examined associations with risk of breast cancer-specific mortality in a total of 96,661 breast cancer patients with survival time data. We then investigated the potential functional role of the selected variants by predicting possible target genes.


12

## Materials and methods

### Breast Cancer Patient Samples

We included data from twelve datasets (n=96,661) in which multiple breast cancer patient cohorts were genotyped by a variety of arrays providing genome-wide coverage of common variants. An overview of the datasets with specification of the arrays used is given in **Supplementary Table 1**. Data from eight of these datasets have been used in previous analyses (n=37,954)[44]. However, the COGS dataset from the Breast Cancer Association Consortium (BCAC) was updated to include additional follow-up and death events and additional genotype data, increasing the number of events and samples to a total of n=29,959 patients. Two new datasets, the BCAC OncoArray and the SUCCESS A trial, comprising 58,027 samples, were added for the current analyses.

The OncoArray is a custom Illumina genotyping array designed by the Genetic Associations and Mechanisms in Oncology (GAME-ON) consortium. It includes 533,000 variants of which 260,660 form a GWAS backbone, with the remainder being custom content, details of which have been described previously[45]. The SUCCESS-A Study[47] is a randomized phase III study of n=3,299 breast cancer cases. Cases from the trial were genotyped using the Illumina Human OmniExpress array. We downloaded imputed genotypes from dbGaP (data reference 6266).

COGS samples that were also genotyped on the OncoArray were removed from the COGS dataset (n=14,426). Female patients with invasive breast cancer diagnosed at age>18 years, and with follow-up data available were included in the analyses. BCAC data from freeze 8 was used, in which 873 COGS samples with unknown breast cancer-specific mortality status were excluded from the analyses. All stages of cancer, including metastatic, were used in


13

the analysis. Some individual studies applied additional selection criteria such as young age or early breast cancer stage (**Supplementary Table 2**).

**Genotype and sample quality control, ancestry analysis and imputation**

The genotype and sample quality control for the datasets have been described previously[44,45,47,48]. Ancestry outliers for each dataset were identified by multidimensional scaling or LAMP[49] on the basis of a set of unlinked variants and HapMap2 populations. Samples of European ancestry were retained for analyses.

Ten of the datasets were imputed using the reference panel from the 1000 Genomes Project in a two-stage procedure. The 1000 Genomes project Phase 3 (October 2014) release was used as the reference panel for all the datasets apart from SUCCESS-A, which used the Phase 1 release (March 2012). Imputation for CGEMS and BPC3 was performed using the program MACH[50]. Phased genotypes were first derived using SHAPEIT[51] and IMPUTE2[52] and then used to perform imputation on the phased data. The main analyses were based on variants that were imputed with imputation $r^2$>0.3 and had minor allele frequency (MAF)>0.01 in at least one of the datasets leading to ~10.4 million variants. To match the individual datasets in the meta-analysis we used the chromosome position. Variants were kept in the analysis as long as they were present in one of the studies. In those cases where there was ambiguity over the naming of the insertions and deletions, the MAF was used for further matching.

**Statistical and bioinformatic methods**

Time-to-event was calculated from the date of diagnosis. For prevalent cases with study entry after diagnosis left truncation was applied, i.e., follow-up started at the date of study


14

entry[53]. Follow-up was right censored on the date of death, on the date last known alive if death did not occur, or at 15 years after diagnosis, whichever came first. We chose the 15 years cut-off because follow-up varied between studies and after that period follow-up data became scarce. Follow-up of the cohorts is illustrated in Kaplan Meier curves (Supplementary Figure 1).

The hazard ratios (HR) for the association of genotypes with breast cancer-specific mortality were estimated using Cox proportional hazards regression implemented in an in-house program written in C++. Analysis of the CGEMS and BPC3 data was conducted using ProbABEL[54]. The estimates of the individual studies were combined using an inverse-variance weighted meta-analysis. Since meta-analysis results based on the Wald test have been shown to be inflated for rare variants[55] we recomputed the standard errors (SE) based on the likelihood ratio test (LRT) statistic (see details in **Supplementary methods)**, using the formula:

$$SE = \log(HR)/\sqrt{LRT}$$

For each dataset we included as covariates a variable number of principal components **(Supplementary Table 1)** from the ancestry analysis as covariates in order to control for cryptic population substructure. The Cox models were stratified by country for the OncoArray dataset and by study for the COGS dataset. Statistical tests were performed for each variant by combining the results for all the datasets using a fixed-effects meta-analysis. Inflation of the test statistics (λ) was estimated by dividing the 45th percentile of the test statistic by 0.357 (the 45th percentile for a $\chi 2$ distribution on 1 degree of freedom). Analyses were carried out for all invasive breast cancer and for ER-positive and ER-negative disease separately.

To assess the probability of a variant being a false positive we used a Bayesian False Discovery Probability (BFDP)[56] test based on the p-value, a prior set to 0.0001 and an upper


15

likely hazard ratio of 1.3.

To predict potential target genes, we used Bedtools v2.26 to intersect notable variants with genomic annotation data relevant to gene regulation activity in samples derived from breast tissue. We examined features including enhancers, promoters and transcription factor binding sites identified by the Roadmap[57] and ENCODE[58] Projects. Expression quantitative loci (eQTL) data from GTEx[59] were queried for evidence of potential cis-regulatory activity.

## Results

Genotype data from 96,661 breast cancer cases (64,171 ER-positive and 16,172 ER-negative) with 7,697 breast cancer deaths within 15 years were included in the primary analyses. For 16,318 cases we did not have ER-status information. The average follow-up time was 6.38 years. Details of the numbers of samples and events in each dataset are given in **Supplementary Table 3**. Manhattan and Quantile-quantile (Q-Q) plots for the associations between variants and breast cancer-specific mortality of all invasive, ER-negative, and ER-positive breast cancers are shown in **Figure 1** and **Figure 2** respectively. There was some evidence of inflation of the test statistic with an inflation factor of 1.06 for all invasive and ER-positive, and 1.05 for ER-negative including all variants. These Q-Q plots showed no evidence of an association at $P<5x10^{-8}$; at less stringent thresholds for significance, there were an increasing number of observed associations for all three analyses **(Figure 2).**

We identified three variants at BFDP<15% associated with breast cancer-specific mortality of patients with ER-negative disease **(Table 1)**. These variants are part of an independent set of 32 highly correlated variants[60] on chromosome 7q21.1 that were associated at $P<5x10^{-6}$ **(Supplementary Table 4)**. The LD matrix between these variants computed based on the 1,000 European genomes[61], and their chromosomal positions, are shown in


16

**Supplementary Figure 2.** The strongest association was for rs67918676: HR=1.27; 95% CI=1.16-1.39; P=1.38x10$^{-7}$; risk allele A frequency=0.12 and BFDP=11%. The imputation efficiency for this variant was high, with $r^2$>0.99 for all datasets.

The lead variant rs67918676 is located in an intron of a long intergenic non-coding RNA gene, *LOC105375207* (AC004009.3), in close proximity to the *HOXA* gene cluster and the lncRNA *HOTTIP*. We tested the genes within a 500 MBp window around the 32 highly correlated variants for the association of their mRNA expression in breast tumours with recurrence-free survival using KMplotter (kmplot.com/analysis). Four of the ten closest genes with probes available showed moderate association with breast cancer survival at P<0.005 (*HOXA9*, *HOTTIP*, *EVX1*, *TAX1BP1*), with these associations mainly observed for ER-negative breast cancer (**Supplementary Table 5A**). Yet, intersecting the germline variants with several sources of genomic annotation information (e.g., chromosome conformation, enhancer-promoter correlations or gene expression) we could not find strong *in silico* evidence of gene regulation by the region containing the associated variants.

We also identified four variants at a BFDP<15% associated with breast cancer-specific mortality of patients with ER-positive disease **(Table 1)**. These variants were part of an independent set of 45 highly correlated variants on chromosome 7q11.22 that were associated at P<5x10$^{-6}$ **(Supplementary Table 6)**. The LD matrix between these variants computed based on the 1,000 European genomes[61], and their chromosomal positions, are shown in **Supplementary Figure 3.** The strongest association was for rs4717568: HR=0.88; 95% CI=0.84; 0.92; P=1.28x10$^{-7}$; risk allele A frequency=0.62 and BFDP=7%. The imputation efficiency for this variant was high, with an average $r^2$=0.96 for all datasets. Two coding genes, *AUTS2* and *GALNT17*, were located within a 500 MBp window around the 45 highly correlated variants, but the expression of neither of the two was associated with breast cancer survival in KMplotter analyses of TCGA data (**Supplementary Table 5B**).


17

The association of rs67918676 with ER-negative breast cancer was observed in eight of nine studies with no significant heterogeneity present at P<0.01 **(Figure 3 and Supplementary Figure 4a)**. For ER-positive disease, the association of rs4717568 was detected in all seven studies with no heterogeneity present at P<0.01 **(Figure 4 and Supplementary Figure 4b)**.

Apart from the 7q variants, only one isolated rare variant reached BFDP values below 15% for all tumours **(Table 1)**. The variant, rs370332736: HR=1.17; 95% CI=1.10; 1.24; $P=2.48x10^{-7}$; risk allele A frequency=0.09 and BFDP=13%, is located on chromosome 6 and has an average imputation efficiency of $r^2$=0.96 for all datasets. In addition, there were several variants found at $P<10^{-6}$ for all three analyses **(Supplementary Table 4, Supplementary Table 6** and **Supplementary Table 7)**.

## Discussion

In this large survival analysis, we report a genome-wide study for identifying genetic markers associated with breast cancer-specific mortality, involving 96,661 patients from a combined meta-analysis. We found one noteworthy region with 32 highly correlated variants on chromosome 7q21.1 for ER-negative. The lead variant rs67918676 ($P=1.38x10^{-7}$ and BFDP of 11% under reasonable assumptions for the prior probability of association) is located in a long intergenic non-coding RNA gene (AC004009.3). While this represents an uncharacterised transcript mainly expressed in testis and prostate, it is located about 200 kb away from a cluster of *HOXA* homeobox genes that has been implicated in breast cancer aetiology and prognosis[62,63]. This region also contains *HOTTIF*, a lncRNA with prognostic value on clinical outcome in breast cancer[64]. The flanking region on the opposite side contains *TAX1BP1*, a gene that may be involved in chemosensitivity[65]. Interestingly, database mining using KMplotter revealed evidence for an association of the expression


18

of these nearby genes with survival from ER-negative breast cancer. On the other hand, the enhancer activity at this noteworthy locus was predicted to be low based on the intersection with biofeatures characteristic of regulatory activity as no known eQTLs appear to exist in this region, suggesting that gene regulatory effects of the identified variants are limited in breast tissue or may be activated under certain untested conditions. For ER-positive tumours, we found another noteworthy region with 45 highly correlated variants at $p<5x10E^{-6}$ on chromosome 7q11.22. The lead variant rs4717568 ($P=1.28x10^{-7}$ and BFDP of 7%) is located between the *AUTS2* and the *GALNT17* genes. *GALNT17* encodes an N-acetylgalactosaminyltransferase that may play a role in membrane trafficking. *AUTS2* has been implicated in neurodevelopment, but *AUTS2* overexpression in cancer has also been linked with resistance to chemotherapy and epithelial-to-mesenchymal transition[69]. It has been postulated that overexpression of *AUTS2* is specific for metastases[69], which may be consistent with the inconspicuous gene expression results in the TCGA database.

It is important to note the differences between the present and the previous GWAS study we had undertaken[44], the latter done in a much smaller dataset (3,632 events versus 7,697 events in the current study) that did not include the OncoArray study. The OncoArray study is the largest dataset used in the present meta-analysis and also the study with the highest imputation quality. The two previously reported variants (rs148760487 for all breast cancer tumours and rs2059614 for ER-negative tumours) were not associated with breast cancer-specific mortality in the current analyses ($P=1.59x10^{-3}$ and $P=5.41x10^{-4}$ respectively). The most likely explanation for this is that the original results were false positive findings, despite the original association being nominally “genome-wide significant”. The BDFPs for the original reported associations were 54% and 16%, respectively. For the lead variants identified in the present analysis, we tested for differences in the imputation quality between the current and previous analysis. All variants had high imputation quality (~0.99) in the previous study,


19

suggesting that the longer and more complete follow-up together with a higher number of events allowed more robust identification of breast cancer mortality associations. However, there are some weaknesses of the current meta-analysis such as heterogeneity between patient treatment over time and between countries and between datasets with different study designs that should be considered. These limitations, intrinsic to large survival meta-analyses, increase the noise and reduce the power to detect true associations.

In conclusion, we found two novel candidate regions at chromosome 7 for breast cancer survival, credible at a BFDP<15% and associated with either ER-negative or ER-positive breast cancer-specific mortality. Concerning additional variants, we might still be underpowered to obtain a more comprehensive picture of genomic markers for breast cancer outcome. Overall, the role of germline variants in breast cancer mortality is still unclear[36,37,66] and additional analyses with larger sample sizes and more complete follow-up including treatments are needed. In addition, alternative methods that integrate multiple data sources such as gene expression, protein-protein interactions or pathway analyses may be used to aggregate the effect of multiple variants with small effects[67]. Such approaches could increase the power of the analyses while better explaining the underlying biological mechanisms associated with breast cancer mortality.

**ACKNOWLEDGEMENTS**

**BCAC:** We thank all the individuals who took part in these studies and all the researchers, clinicians, technicians and administrative staff who have enabled this work to be carried out. We acknowledge all contributors to the COGS and OncoArray study design, chip design, genotyping, and genotype analyses. **ABCFS** thank Maggie Angelakos, Judi Maskiell, Gillian Dite. **ABCS** thanks the Blood bank Sanquin, The Netherlands. **ABCTB** Investigators: Christine Clarke,Rosemary Balleine, Robert Baxter,Stephen Braye, Jane Carpenter, Jane Dahlstrom, John Forbes, Soon Lee, Debbie Marsh, Adrienne Morey, Nirmala Pathmanathan, Rodney Scott, Allan Spigelman, Nicholas Wilcken, Desmond Yip. Samples are made available to researchers on a non-exclusive basis. **BBCS**


20

685thanks Eileen Williams, Elaine Ryder-Mills, Kara Sargus. The **BCINIS** study would not have been possible
686without the contributions of Dr. K. Landsman, Dr. N. Gronich, Dr. A. Flugelman, Dr. W. Saliba, Dr. E. Liani, Dr. I.
687Cohen, Dr. S. Kalet, Dr. V. Friedman, Dr. O. Barnet of the NICCC in Haifa, and all the contributing family medi-
688cine, surgery, pathology and oncology teams in all medical institutes in Northern Israel. **BIGGS** thanks Niall McIn-
689erney, Gabrielle Colleran, Andrew Rowan, Angela Jones. The **BREOGAN** study would not have been possible
690without the contributions of the following: Manuela Gago-Dominguez, Jose Esteban Castelao, Angel Carracedo,
691Victor Muñoz Garzón, Alejandro Novo Domínguez, Maria Elena Martinez, Sara Miranda Ponte, Carmen Redondo
692Marey, Maite Peña Fernández, Manuel Enguix Castelo, Maria Torres, Manuel Calaza (BREOGAN), José An-
693túnez, Máximo Fraga and the staff of the Department of Pathology and Biobank of the University Hospital Com-
694plex of Santiago-CHUS, Instituto de Investigación Sanitaria de Santiago, IDIS, Xerencia de Xestion Integrada de
695Santiago-SERGAS; Joaquín González-Carreró and the staff of the Department of Pathology and Biobank of Uni-
696versity Hospital Complex of Vigo, Instituto de Investigacion Biomedica Galicia Sur, SERGAS, Vigo, Spain.
697**BSUCH** thanks Peter Bugert, Medical Faculty Mannheim. **CCGP** thanks Styliani Apostolaki, Anna Margiolaki,
698Georgios Nintos, Maria Perraki, Georgia Saloustrou, Georgia Sevastaki, Konstantinos Pompodakis. **CGPS**
699thanks staff and participants of the Copenhagen General Population Study. For the excellent technical assist-
700ance: Dorthe Uldall Andersen, Maria Birna Arnadottir, Anne Bank, Dorthe Kjeldgård Hansen. The Danish Cancer
701Biobank is acknowledged for providing infrastructure for the collection of blood samples for the cases. **CNIO-BCS**
702thanks Guillermo Pita, Charo Alonso, Nuria Álvarez, Pilar Zamora, Primitiva Menendez, the Human Genotyping-
703CEGEN Unit (CNIO). Investigators from the **CPS-II** cohort thank the participants and Study Management Group
704for their invaluable contributions to this research. They also acknowledge the contribution to this study from cent-
705ral cancer registries supported through the Centers for Disease Control and Prevention National Program of Can-
706cer Registries, as well as cancer registries supported by the National Cancer Institute Surveillance Epidemiology
707and End Results program. The **CTS** Steering Committee includes Leslie Bernstein, Susan Neuhausen, James
708Lacey, Sophia Wang, Huiyan Ma, and Jessica Clague DeHart at the Beckman Research Institute of City of
709Hope, Dennis Deapen, Rich Pinder, and Eunjung Lee at the University of Southern California, Pam Horn-Ross,
710Peggy Reynolds, Christina Clarke Dur and David Nelson at the Cancer Prevention Institute of California, Hoda
711Anton-Culver, Argyrios Ziogas, and Hannah Park at the University of California Irvine, and Fred Schumacher at
712Case Western University. **DIETCOMPLYF** thanks the patients, nurses and clinical staff involved in the study. The
713DietCompLyf study was funded by the charity Against Breast Cancer (Registered Charity Number 1121258) and
714the NCRN. We thank the participants and the investigators of EPIC (European Prospective Investigation into
715Cancer and Nutrition). **ESTHER** thanks Hartwig Ziegler, Sonja Wolf, Volker Hermann, Christa Stegmaier, Katja

Butterbach. **FHRISK** thanks NIHR for funding. **GC-HBOC** thanks Stefanie Engert, Heide Hellebrand, Sandra Kröber and LIFE - Leipzig Research Centre for Civilization Diseases (Markus Loeffler, Joachim Thiery, Matthias Nüchter, Ronny Baber). The **GENICA** Network: Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, and University of Tübingen, Germany [HB, WYL], German Cancer Consortium (DKTK) and German Cancer Research Center (DKFZ) [HB], Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany [YDK, Christian Baisch], Institute of Pathology, University of Bonn, Germany [Hans-Peter Fischer], Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany [UH], Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, Germany [Thomas Brüning, Beate Pesch, Sylvia Rabstein, Anne Lotz]; and Institute of Occupational Medicine and Maritime Medicine, University Medical Center Hamburg-Eppendorf, Germany [Volker Harth]. **HABCS** thanks Michael Bremer. **HEBCS** thanks, Rainer Fagerholm, Kirsimari Aaltonen, Karl von Smitten, Irja Erkkilä. **HUBCS** thanks Shamil Gantsev. **KARMA** and **SASBAC** thank the Swedish Medical Research Counsel. **KBCP** thanks Eija Myöhänen, Helena Kemiläinen. **kConFab**/**AOCS** wish to thank Heather Thorne, Eveline Niedermayr, all the kConFab research nurses and staff, the heads and staff of the Family Cancer Clinics, and the Clinical Follow Up Study (which has received funding from the NHMRC, the National Breast Cancer Foundation, Cancer Australia, and the National Institute of Health (USA)) for their contributions to this resource, and the many families who contribute to kConFab. **LMBC** thanks Gilian Peuteman, Thomas Van Brussel, EvyVanderheyden and Kathleen Corthouts. **MARIE** thanks Petra Seibold, Judith Heinz, Nadia Obi, Alina Vrieling, Sabine Behrens, Ursula Eilber, Muhabbet Celik, Til Olchers and Stefan Nickels. **MBCSG**: Paolo Peterlongo, Bernard Peissel, Roberto Villa, Cristina Zanzottera, Irene Feroce, and the personnel of the Cogentech Cancer Genetic Test Laboratory. We thank the coordinators, the research staff and especially the MMHS participants for their continued collaboration on research studies in breast cancer. The following are NBCS Collaborators: Kristine K. Sahlberg (PhD), Lars Ottestad (MD), Rolf Kåresen (Prof. Em.) Dr. Ellen Schlichting (MD), Marit Muri Holmen (MD), Toril Sauer (MD), Vilde Haakensen (MD), Olav Engebråten (MD), Bjørn Naume (MD), Alexander Fosså (MD), Cecile E. Kiserud (MD), Kristin V. Reinertsen (MD), Åslaug Helland (MD), Margit Riis (MD), Jürgen Geisler (MD) and OSBREAC. **NHS**/**NHS2** would like to thank the participants and staff of the NHS and NHS2 for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. **OBCS** thanks Arja Jukkola-Vuorinen, Mervi Grip, Saila Kauppila, Meeri Otsukka, Leena Keskitalo and Kari Mononen for their contributions to this study. **OFBCR** thanks Teresa Selander, Nayana Weerasooriya. **ORIGO** thanks E. Krol-Warmerdam, and J. Blom for patient ac-

crual, administering questionnaires, and managing clinical information. **PBCS** thanks Louise Brinton, Mark Sherman, Neonila Szeszenia-Dabrowska, Beata Peplonska, Witold Zatonski, Pei Chao, Michael Stagner. The ethical approval for the **POSH** study is MREC /00/6/69, UKCRN ID: 1137. We thank staff in the Experimental Cancer Medicine Centre (ECMC) supported Faculty of Medicine Tissue Bank and the Faculty of Medicine DNA Banking resource. **PREFACE** thanks Sonja Oeser and Silke Landrith. **PROCAS** thanks NIHR for funding. **RBCS** thanks Petra Bos, Jannet Blom, Ellen Crepin, Elisabeth Huijskens, Anja Kromwijk-Nieuwlaat, Annette Heemskerk, the Erasmus MC Family Cancer Clinic. **SBCS** thanks Sue Higham, Helen Cramp, Dan Connley, Ian Brock, Sabapathy Balasubramanian and Malcolm W.R. Reed. We thank the **SEARCH** and **EPIC** teams. **SKKDKFZS** thanks all study participants, clinicians, family doctors, researchers and technicians for their contributions and commitment to this study. We thank the SUCCESS Study teams in Munich, Duessldorf, Erlangen and Ulm. We thank the **SUCCESS** Study teams in Munich, Duessldorf, Erlangen and Ulm. **SZBCS** thanks Ewa Putresza. **UCIBCS** thanks Irene Masunaka. **UKBGS** thanks Breast Cancer Now and the Institute of Cancer Research for support and funding of the Breakthrough Generations Study, and the study participants, study staff, and the doctors, nurses and other health care providers and health information sources who have contributed to the study. We acknowledge NHS funding to the Royal Marsden/ICR NIHR Biomedical Research Centre. The authors thank the **WHI** investigators and staff for their dedication and the study participants for making the program possible.

## AUTHOR CONTRIBUTIONS

M.K.S. and P.D.P.F. conceived the study. Q.G., M.E.G., S.K., C.J.T. and T.D. performed the data analyses. M.K.S., P.D.P.F., Q.G., M.E.G., T.D. and D.M.E. were involved in the interpretation of the data. J.D., D.F.E., P.D.P.F. , S.C. and J.B. provided statistical and computational support for the data analyses. R.K, Q.W., M.K.B. and J.D. provided database support. M.E.G., Q.G., T.D., M.K.S. and P.D.P.F. wrote the first draft of the manuscript. All authors contributed data from their own studies, helped revise the manuscript, and approved the final version.

## ADDITIONAL INFORMATION

**Funding:** **BCAC** is funded by Cancer Research UK [C1287/A16563, C1287/A10118], the European Union's Horizon 2020 Research and Innovation Programme (grant numbers 634935 and 633784 for BRIDGES and B-CAST respectively), and by the European Community´s Seventh Framework Programme under grant agreement number 223175 (grant number HEALTH-F2-2009-223175) (COGS). The EU Horizon 2020 Research and Innovation


23

777Programme funding source had no role in study design, data collection, data analysis, data interpretation or writ-
778ing of the report.

779Genotyping of the **OncoArray** was funded by the NIH Grant U19 CA148065, and Cancer UK Grant
780C1287/A16563 and the PERSPECTIVE project supported by the Government of Canada through Genome
781Canada and the Canadian Institutes of Health Research (grant GPH-129344) and, the Ministère de l'Économie,
782Science et Innovation du Québec through Genome Québec and the PSRSIIRI-701 grant, and the Quebec Breast
783Cancer Foundation. Funding for the **iCOGS** infrastructure came from: the European Community's Seventh
784Framework Programme under grant agreement n° 223175 (HEALTH-F2-2009-223175) (COGS), Cancer Re-
785search UK (C1287/A10118, C1287/A10710, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007,
786C5047/A10692, C8197/A16565), the National Institutes of Health (CA128978) and Post-Cancer GWAS initiative
787(1U19 CA148537, 1U19 CA148065 and 1U19 CA148112 - the GAME-ON initiative), the Department of Defence
788(W81XWH-10-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks
789of Breast Cancer, and Komen Foundation for the Cure, the Breast Cancer Research Foundation, and the Ovari-
790an Cancer Research Fund. The DRIVE Consortium was funded by U19 CA148065.

791**ABCFS** was supported by grant UM1 CA164920 from the National Cancer Institute (USA). The content of this
792manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collab-
793orating centres in the in the Breast Cancer Family Registry (**BCFR)**, nor does mention of trade names, commer-
794cial products, or organizations imply endorsement by the USA Government or the BCFR. The **ABCFS** was also
795supported by the National Health and Medical Research Council of Australia, the New South Wales Cancer
796Council, the Victorian Health Promotion Foundation (Australia) and the Victorian Breast Cancer Research Con-
797sortium. J.L.H. is a National Health and Medical Research Council (NHMRC) Senior Principal Research Fel-
798low. M.C.S. is a NHMRC Senior Research Fellow. The **ABCS** study was supported by the Dutch Cancer Society
799[grants NKI 2007-3839; 2009-4363; 2015-7632]. The **ABCTB** is generously supported by the National Health and
800Medical Research Council of Australia, The Cancer Institute NSW and the National Breast Cancer Foundation.
801The work of the **BBCC** was partly funded by ELAN-Fond of the University Hospital of Erlangen. The **BBCS** is fun-
802ded by Cancer Research UK and Breast Cancer Now and acknowledges NHS funding to the NIHR Biomedical
803Research Centre, and the National Cancer Research Network (NCRN). For the **BCFR-NY, BCFR-PA, BCFR-UT**
804this work was supported by grant UM1 CA164920 from the National Cancer Institute. For **BIGGS**, ES is suppor-
805ted by NIHR Comprehensive Biomedical Research Centre, Guy's & St. Thomas' NHS Foundation Trust in part-
806nership with King's College London, United Kingdom. IT is supported by the Oxford Biomedical Research Centre.
807The **BREOGAN** is funded by Acción Estratégica de Salud del Instituto de Salud Carlos III FIS PI12/02125/Cofi-

808nanciado FEDER; Acción Estratégica de Salud del Instituto de Salud Carlos III FIS Intrasalud (PI13/01136); Pro-
809grama Grupos Emergentes, Cancer Genetics Unit, Instituto de Investigacion Biomedica Galicia Sur. Xerencia de
810Xestion Integrada de Vigo-SERGAS, Instituto de Salud Carlos III, Spain; Grant 10CSA012E, Consellería de In-
811dustria Programa Sectorial de Investigación Aplicada, PEME I + D e I + D Suma del Plan Gallego de Investiga-
812ción, Desarrollo e Innovación Tecnológica de la Consellería de Industria de la Xunta de Galicia, Spain; Grant
813EC11-192. Fomento de la Investigación Clínica Independiente, Ministerio de Sanidad, Servicios Sociales e Igual-
814dad, Spain; and Grant FEDER-Innterconecta. Ministerio de Economia y Competitividad, Xunta de Galicia, Spain.
815The **BSUCH** study was supported by the Dietmar-Hopp Foundation, the Helmholtz Society and the German Can-
816cer Research Center (DKFZ). **CCGP** is supported by funding from the University of Crete. The **CECILE** study
817was supported by Fondation de France, Institut National du Cancer (INCa), Ligue Nationale contre le Cancer,
818Agence Nationale de Sécurité Sanitaire, de l'Alimentation, de l'Environnement et du Travail (ANSES), Agence
819Nationale de la Recherche (ANR). The **CGPS** was supported by the Chief Physician Johan Boserup and Lise
820Boserup Fund, the Danish Medical Research Council, and Herlev and Gentofte Hospital. The **CNIO-BCS** was
821supported by the Instituto de Salud Carlos III, the Red Temática de Investigación Cooperativa en Cáncer and
822grants from the Asociación Española Contra el Cáncer and the Fondo de Investigación Sanitario (PI11/00923
823and PI12/00070). The American Cancer Society funds the creation, maintenance, and updating of the **CPS-II** co-
824hort. The **CTS** was initially supported by the California Breast Cancer Act of 1993 and the California Breast Can-
825cer Research Fund (contract 97-10500) and is currently funded through the National Institutes of Health (R01
826CA77398, UM1 CA164917, and U01 CA199277). Collection of cancer incidence data was supported by the Cali-
827fornia Department of Public Health as part of the statewide cancer reporting program mandated by California
828Health and Safety Code Section 103885. The University of Westminster curates the **DietCompLyf** database fun-
829ded by Against Breast Cancer Registered Charity No. 1121258 and the NCRN. The coordination of **EPIC** is finan-
830cially supported by the European Commission (DG-SANCO) and the International Agency for Research on Can-
831cer. The national cohorts are supported by: Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale
832de l'Education Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM) (France); German
833Cancer Aid, German Cancer Research Center (DKFZ), Federal Ministry of Education and Research (BMBF)
834(Germany); the Hellenic Health Foundation, the Stavros Niarchos Foundation (Greece); Associazione Italiana per
835la Ricerca sul Cancro-AIRC-Italy and National Research Council (Italy); Dutch Ministry of Public Health, Welfare
836and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch
837ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (The Nether-
838lands); Health Research Fund (FIS), PI13/00061 to Granada, PI13/01162 to EPIC-Murcia, Regional Govern-


25

ments of Andalucía, Asturias, Basque Country, Murcia and Navarra, ISCIII RETIC (RD06/0020) (Spain); Cancer Research UK (14136 to EPIC-Norfolk; C570/A16491 and C8221/A19170 to EPIC-Oxford), Medical Research Council (1000143 to EPIC-Norfolk, MR/M012190/1 to EPIC-Oxford) (United Kingdom). The **ESTHER** study was supported by a grant from the Baden Württemberg Ministry of Science, Research and Arts. Additional cases were recruited in the context of the VERDI study, which was supported by a grant from the German Cancer Aid (Deutsche Krebshilfe). **FHRISK** is funded from NIHR grant PGfAR 0707-10031. The **GC-HBOC** is supported by the German Cancer Aid (grant no 110837, coordinator: Rita K. Schmutzler, Cologne). This work was also funded by the European Regional Development Fund and Free State of Saxony, Germany (LIFE - Leipzig Research Centre for Civilization Diseases, project numbers 713-241202, 713-241202, 14505/2470, 14575/2470). The **GENICA** was funded by the Federal Ministry of Education and Research (BMBF) Germany grants 01KW9975/5, 01KW9976/8, 01KW9977/0 and 01KW0114, the Robert Bosch Foundation, Stuttgart, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, the Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, as well as the Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany. The **GESBC** was supported by the Deutsche Krebshilfe e. V. [70492] and the German Cancer Research Center (DKFZ). The **HABCS** study was supported by the Claudia von Schilling Foundation for Breast Cancer Research, by the Lower Saxonian Cancer Society, and by the Rudolf Bartling Foundation. The **HEBCS** was financially supported by the Helsinki University Central Hospital Research Fund, Academy of Finland (266528), the Finnish Cancer Society, and the Sigrid Juselius Foundation. The **HUBCS** was supported by a grant from the German Federal Ministry of Research and Education (RUS08/017), and by the Russian Foundation for Basic Research and the Federal Agency for Scientific Organizations for support the Bioresource collections and RFBR grants 14-04-97088, 17-29-06014 and 17-44-020498. Financial support for **KARBAC** was provided through the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and Karolinska Institutet, the Swedish Cancer Society, The Gustav V Jubilee foundation and Bert von Kantzows foundation. The **KARMA** study was supported by Märit and Hans Rausings Initiative Against Breast Cancer. The **KBCP** was financially supported by the special Government Funding (EVO) of Kuopio University Hospital grants, Cancer Fund of North Savo, the Finnish Cancer Organizations, and by the strategic funding of the University of Eastern Finland. **kConFab** is supported by a grant from the National Breast Cancer Foundation, and previously by the National Health and Medical Research Council (NHMRC), the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania and South Australia, and the Cancer Foundation of Western Australia. **LMBC** is supported by the 'Stichting tegen Kanker'. The **MARIE** study was supported by the Deutsche Krebshilfe e.V. [70-2892-BR I,


26

870106332, 108253, 108419, 110826, 110828], the Hamburg Cancer Society, the German Cancer Research Center
871(DKFZ) and the Federal Ministry of Education and Research (BMBF) Germany [01KH0402]. **MBCSG** is suppor-
872ted by grants from the Italian Association for Cancer Research (AIRC) and by funds from the Italian citizens who
873allocated the 5/1000 share of their tax payment in support of the Fondazione IRCCS Istituto Nazionale Tumori,
874according to Italian laws (INT-Institutional strategic projects "5x1000"). The **MCBCS** was supported by the NIH
875grants CA192393, CA116167, CA176785 an NIH Specialized Program of Research Excellence (SPORE) in
876Breast Cancer [CA116201], and the Breast Cancer Research Foundation and a generous gift from the David F.
877and Margaret T. Grohne Family Foundation. **MCCS** cohort recruitment was funded by VicHealth and Cancer
878Council Victoria. The **MCCS** was further supported by Australian NHMRC grants 209057 and 396414, and by in-
879frastructure provided by Cancer Council Victoria. Cases and their vital status were ascertained through the Vic-
880torian Cancer Registry (VCR) and the Australian Institute of Health and Welfare (AIHW), including the National
881Death Index and the Australian Cancer Database. The **MEC** was supported by NIH grants CA63464, CA54281,
882CA098758, CA132839 and CA164973. The **MISS** study is supported by funding from ERC-2011-294576 Ad-
883vanced grant, Swedish Cancer Society, Swedish Research Council, Local hospital funds, Berta Kamprad Found-
884ation, Gunnar Nilsson. The **MMHS** study was supported by NIH grants CA97396, CA128931, CA116201,
885CA140286 and CA177150. The work of **MTLGEBCS** was supported by the Quebec Breast Cancer Foundation,
886the Canadian Institutes of Health Research for the "CIHR Team in Familial Risks of Breast Cancer" program –
887grant # CRN-87521 and the Ministry of Economic Development, Innovation and Export Trade – grant # PSR-
888SIIRI-701. The **NBCS** has received funding from the K.G. Jebsen Centre for Breast Cancer Research; the Re-
889search Council of Norway grant 193387/V50 (to A-L Børresen-Dale and V.N. Kristensen) and grant 193387/H10
890(to A-L Børresen-Dale and V.N. Kristensen), South Eastern Norway Health Authority (grant 39346 to A-L Bør-
891resen-Dale) and the Norwegian Cancer Society (to A-L Børresen-Dale and V.N. Kristensen). The **NC-BCFR** and
892**OFBCR** were supported by grant UM1 CA164920 from the National Cancer Institute (USA). The **NCBCS** was
893funded by Komen Foundation, the National Cancer Institute (P50 CA058223, U54 CA156733, U01 CA179715),
894and the North Carolina University Cancer Research Fund. The **NHS** was supported by NIH grants P01 CA87969,
895UM1 CA186107, and U19 CA148065. The **NHS2** was supported by NIH grants UM1 CA176726 and U19
896CA148065. The **OBCS** was supported by research grants from the Finnish Cancer Foundation, the Academy of
897Finland (grant number 250083, 122715 and Center of Excellence grant number 251314), the Finnish Cancer
898Foundation, the Sigrid Juselius Foundation, the University of Oulu, the University of Oulu Support Foundation
899and the special Governmental EVO funds for Oulu University Hospital-based research activities. The **ORIGO**
900study was supported by the Dutch Cancer Society (RUL 1997-1505) and the Biobanking and Biomolecular Re-


27

sources Research Infrastructure (BBMRI-NL CP16). The **PBCS** was funded by Intramural Research Funds of the National Cancer Institute, Department of Health and Human Services, USA. Genotyping for **PLCO** was supported by the Intramural Research Program of the National Institutes of Health, NCI, Division of Cancer Epidemiology and Genetics. The **PLCO** is supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics and supported by contracts from the Division of Cancer Prevention, National Cancer Institute, National Institutes of Health. The **POSH** study is funded by Cancer Research UK (grants C1275/A11699, C1275/C22524, C1275/A19187, C1275/A15956 and Breast Cancer Campaign 2010PR62, 2013PR044. PROCAS is funded from NIHR grant PGfAR 0707-10031. **PROCAS** is funded from NIHR grant PGfAR 0707-10031. The **RBCS** was funded by the Dutch Cancer Society (DDHK 2004-3124, DDHK 2009-4318). The **SASBAC** study was supported by funding from the Agency for Science, Technology and Research of Singapore (A*STAR), the US National Institute of Health (NIH) and the Susan G. Komen Breast Cancer Foundation. The **SBCS** was supported by Sheffield Experimental Cancer Medicine Centre and Breast Cancer Now Tissue Bank. **SEARCH** is funded by Cancer Research UK [C490/A10124, C490/A16561] and supported by the UK National Institute for Health Research Biomedical Research Centre at the University of Cambridge. The University of Cambridge has received salary support for PDPP from the NHS in the East of England through the Clinical Academic Reserve. **SKKDKFZS** is supported by the DKFZ. The **SMC** is funded by the Swedish Cancer Foundation. The **SZBCS** was supported by Grant PBZ_KBN_122/P05/2004. The **UCIBCS** component of this research was supported by the NIH [CA58860, CA92044] and the Lon V Smith Foundation [LVS39420]. The **UKBGS** is funded by Breast Cancer Now and the Institute of Cancer Research (ICR), London. ICR acknowledges NHS funding to the NIHR Biomedical Research Centre. The **USRT** Study was funded by Intramural Research Funds of the National Cancer Institute, Department of Health and Human Services, USA. The **WHI** program is funded by the National Heart, Lung, and Blood Institute, the US National Institutes of Health and the US Department of Health and Human Services (HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C and HHSN271201100004C). This work was also funded by NCI U19 CA148065-01.

**Ethics approval:** The study was performed in accordance with the Declaration of Helsinki. All individual studies, from which data was used, were approved by the appropriate medical ethical committees and/or institutional review boards. All study participants provided informed consent.



28

# References


1. IARC. Available at: http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx.
2. Hartman, M. *et al.* Is breast cancer prognosis inherited? *Breast Cancer Res.* **9,** R39 (2007).
3. Lindström, L. S. *et al.* Familial concordance in cancer survival: a Swedish population-based study. *Lancet Oncol.* **8,** 1001–6 (2007).
4. Udler, M. & Pharoah, P. D. Germline genetic variation and breast cancer survival: prognostic and therapeutic implications. *Futur. Oncol.* **3,** 491–495 (2007).
5. Verkooijen, H. M. *et al.* Breast cancer prognosis is inherited independently of patient, tumor and treatment characteristics. *Int. J. Cancer* **130,** 2103–2110 (2012).
6. Broeks, A. *et al.* Low penetrance breast cancer susceptibility loci are associated with specific breast tumor subtypes: findings from the Breast Cancer Association Consortium. *Hum. Mol. Genet.* **20,** 3289–303 (2011).
7. Yang, X. R. *et al.* Associations of breast cancer risk factors with tumor subtypes: a pooled analysis from the Breast Cancer Association Consortium studies. *JNCI J. Natl. Cancer Inst.* **103,** 250–263 (2011).
8. Blows, F. M. *et al.* Subtyping of Breast Cancer by Immunohistochemistry to Investigate a Relationship between Subtype and Short and Long Term Survival: A Collaborative Analysis of Data for 10,159 Cases from 12 Studies. *PLoS Med.* **7,** e1000279 (2010).
9. Fagerholm, R. *et al.* NAD(P)H:quinone oxidoreductase 1 NQO1*2 genotype (P187S) is a strong prognostic and predictive factor in breast cancer. *Nat. Genet.* **40,** 844–53 (2008).
10. Hoskins, J. M., Carey, L. A. & McLeod, H. L. CYP2D6 and tamoxifen: DNA matters in breast cancer. *Nat. Rev. Cancer* **9,** 576–586 (2009).
11. Koutras, A., Kotoula, V. & Fountzilas, G. Prognostic and predictive role of vascular endothelial growth factor polymorphisms in breast cancer. *Pharmacogenomics* **16,** 79–94 (2015).
12. Hein, A. *et al.* Genetic variants in VEGF pathway genes in neoadjuvant breast cancer patients receiving bevacizumab: Results from the randomized phase III GeparQuinto study. *Int. J. cancer* **137,** 2981–8 (2015).
13. Hsieh, S. M., Lintell, N. a & Hunter, K. W. Germline polymorphisms are potential metastasis risk and prognosis markers in breast cancer. *Breast Dis.* **26,** 157–62 (2007).
14. Crawford, N. P. S. *et al.* Germline polymorphisms in SIPA1 are associated with metastasis and other indicators of poor prognosis in breast cancer. *Breast Cancer Res.* **8,** R16 (2006).
15. Paulsson, J. & Micke, P. Prognostic relevance of cancer-associated fibroblasts in human cancer. *Semin. Cancer Biol.* **25,** 61–8 (2014).
16. Winslow, S., Leandersson, K., Edsjö, A. & Larsson, C. Prognostic stromal gene signatures in breast cancer. *Breast Cancer Res.* **17,** 23 (2015).
17. Loi, S. *et al.* Prognostic and Predictive Value of Tumor-Infiltrating Lymphocytes in a Phase III Randomized Adjuvant Breast Cancer Trial in Node-Positive Breast Cancer Comparing the Addition of Docetaxel to Doxorubicin With Doxorubicin-Based Chemotherapy: BIG 02-98. *J. Clin. Oncol.* **31,** 860–867 (2013).
18. Ali, H. R. *et al.* Association between CD8+ T-cell infiltration and breast cancer survival in 12,439 patients. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **25,** 1536–43 (2014).
19. Ambrosone, C. B. *et al.* Polymorphisms in glutathione S-transferases (GSTM1 and GSTT1) and survival after treatment for breast cancer. *Cancer Res.* **61,** 7130–5 (2001).
20. Goode, E. L. *et al.* Effect of germ-line genetic variation on breast cancer survival in a population-based study. *Cancer Res.* **62,** 3052–7 (2002).
21. Ambrosone, C. B. *et al.* Polymorphisms in genes related to oxidative stress (MPO, MnSOD, CAT) and survival after treatment for breast cancer. *Cancer Res.* **65,** 1105–11 (2005).
22. Boersma, B. J. *et al.* Association of breast cancer outcome with status of p53 and MDM2 SNP309. *J. Natl. Cancer Inst.* **98,** 911–9 (2006).
23. Thussbas, C. *et al.* FGFR4 Arg388 Allele Is Associated With Resistance to Adjuvant Therapy in Primary Breast Cancer. *J. Clin. Oncol.* **24,** 3747–3755 (2006).
24. Decock, J. *et al.* Association of Matrix Metalloproteinase-8 Gene Variation with Breast Cancer Prognosis. *Cancer Res.* **67,** 10214–10221 (2007).
25. Hughes, S. *et al.* Matrix metalloproteinase single-nucleotide polymorphisms and haplotypes predict breast cancer progression. *Clin. Cancer Res.* **13,** 6673–80 (2007).
26. Udler, M. *et al.* Common germline genetic variation in antioxidant defense genes and survival after diagnosis of breast cancer. *J. Clin. Oncol.* **25,** 3015–23 (2007).
27. Einarsdóttir, K. *et al.* ESR1 and EGFgenetic variation in relation to breast cancer risk and survival. *Breast Cancer Res.* **10,** R15 (2008).
28. Fasching, P. A. *et al.* Single nucleotide polymorphisms of the aromatase gene (CYP19A1), HER2/neu



29

status, and prognosis in breast cancer patients. *Breast Cancer Res. Treat.* **112,** 89–98 (2008).
29. Schmidt, M. K. *et al.* Combined effects of single nucleotide polymorphisms TP53 R72P and MDM2 SNP309, and p53 expression on survival of breast cancer patients. *Breast Cancer Res.* **11,** R89 (2009).
30. Varadi, V. *et al.* Polymorphisms in telomere-associated genes, breast cancer susceptibility and prognosis. *Eur. J. Cancer* **45,** 3008–3016 (2009).
31. Lin, W.-Y. *et al.* A role for XRCC2 gene polymorphisms in breast cancer risk and survival. *J. Med. Genet.* **48,** 477–484 (2011).
32. Fasching, P. A. *et al.* The role of genetic breast cancer susceptibility variants as prognostic factors. *Hum. Mol. Genet.* **21,** 3926–39 (2012).
33. Barrdahl, M. *et al.* Association of breast cancer risk loci with breast cancer survival. *Int. J. Cancer* **137,** 2837–2845 (2015).
34. Jamshidi, M. *et al.* SNP-SNP interaction analysis of NF-κB signaling pathway on breast cancer survival. *Oncotarget* **6,** 37979–94 (2015).
35. Weischer, M. *et al.* CHEK2*1100delC heterozygosity in women with breast cancer associated with early death, breast cancer-specific death, and increased risk of a second breast cancer. *J. Clin. Oncol.* **30,** 4308–16 (2012).
36. Pirie, A. *et al.* Common germline polymorphisms associated with breast cancer-specific survival. *Breast Cancer Res.* **17,** 58 (2015).
37. Azzato, E. M. *et al.* A Genome-Wide Association Study of Prognosis in Breast Cancer. *Cancer Epidemiol. Biomarkers Prev.* **19,** 1140–1143 (2010).
38. Azzato, E. M. *et al.* Association between a germline OCA2 polymorphism at chromosome 15q13.1 and estrogen receptor-negative breast cancer survival. *J. Natl. Cancer Inst.* **102,** 650–62 (2010).
39. Kiyotani, K. *et al.* A genome-wide association study identifies locus at 10q22 associated with clinical outcomes of adjuvant tamoxifen therapy for breast cancer patients in Japanese. *Hum. Mol. Genet.* **21,** 1665–72 (2012).
40. Shu, X. O. *et al.* Novel genetic markers of breast cancer survival identified by a genome-wide association study. *Cancer Res.* **72,** 1182–9 (2012).
41. Rafiq, S. *et al.* Identification of inherited genetic variations influencing prognosis in early-onset breast cancer. *Cancer Res.* **73,** 1883–91 (2013).
42. Rafiq, S. *et al.* A genome wide meta-analysis study for identification of common variation associated with breast cancer prognosis. *PLoS One* **9,** e101488 (2014).
43. Michailidou, K. *et al.* Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nat. Genet.* **47,** 373–380 (2015).
44. Guo, Q. *et al.* Identification of novel genetic markers of breast cancer survival. *J. Natl. Cancer Inst.* **107,** djv081-djv081 (2015).
45. Amos, C. I. *et al.* The OncoArray Consortium: A Network for Understanding the Genetic Architecture of Common Cancers. *Cancer Epidemiol. Biomarkers Prev.* **26,** 126–135 (2017).
46. Michailidou, K. *et al.* Association analysis identifies 65 new breast cancer risk loci. *Nature* **551,** 92–94 (2017).
47. dbGaP (SUCCESS). Available at: https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000547.v1.p1.
48. van den Broek, A. J. *et al.* Impact of Age at Primary Breast Cancer on Contralateral Breast Cancer Risk in BRCA1/2 Mutation Carriers. *J. Clin. Oncol.* **34,** 409–18 (2016).
49. Sankararaman, S., Sridhar, S., Kimmel, G. & Halperin, E. Estimating Local Ancestry in Admixed Populations. *Am. J. Hum. Genet.* **82,** 290–303 (2008).
50. Li, Y., Willer, C., Sanna, S. & Abecasis, G. Genotype Imputation. *Annu. Rev. Genomics Hum. Genet.* **10,** 387–406 (2009).
51. Delaneau, O., Marchini, J. & Zagury, J.-F. A linear complexity phasing method for thousands of genomes. *Nat. Methods* **9,** 179–81 (2011).
52. Howie, B., Marchini, J. & Stephens, M. Genotype imputation with thousands of genomes. *G3 (Bethesda).* **1,** 457–70 (2011).
53. Azzato, E. M. *et al.* Prevalent cases in observational studies of cancer survival: do they bias hazard ratio estimates? *Br. J. Cancer* **100,** 1806–1811 (2009).
54. Aulchenko, Y. S., Ripke, S., Isaacs, A. & van Duijn, C. M. GenABEL: an R library for genome-wide association analysis. *Bioinformatics* **23,** 1294–1296 (2007).
55. Ma, C., Blackwell, T., Boehnke, M., Scott, L. J. & GoT2D investigators. Recommended joint and meta-analysis strategies for case-control association testing of single low-count variants. *Genet. Epidemiol.* **37,** 539–50 (2013).
56. Wakefield, J. A Bayesian Measure of the Probability of False Discovery in Genetic Epidemiology Studies. *Am. J. Hum. Genet.* **81,** 208–227 (2007).
57. Roadmap Epigenomics Consortium *et al.* Integrative analysis of 111 reference human epigenomes.


30

*Nature* **518,** 317–329 (2015).
58. Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489,** 57–74 (2012).
59. Aguet, F. *et al.* Genetic effects on gene expression across human tissues. *Nature* **550,** 204–213 (2017).
60. Edwards, S. L., Beesley, J., French, J. D. & Dunning, A. M. Beyond GWASs: illuminating the dark road from association to function. *Am. J. Hum. Genet.* **93,** 779–797 (2013).
61. Machiela, M. J. & Chanock, S. J. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* **31,** 3555–7 (2015).
62. Novak, P. *et al.* Epigenetic Inactivation of the HOXA Gene Cluster in Breast Cancer. *Cancer Res.* **66,** 10664–10670 (2006).
63. Xia, B. *et al.* Homeobox A11 hypermethylation indicates unfavorable prognosis in breast cancer. *Oncotarget* **8,** 9794–9805 (2017).
64. Yang, Y., Qian, J., Xiang, Y., Chen, Y. & Qu, J. The prognostic value of long noncoding RNA HOTTIP on clinical outcomes in breast cancer. *Oncotarget* **8,** 6833–6844 (2017).
65. Choi, H. & Lee, S. K. TAX1BP1 downregulation by EBV-miR-BART15-3p enhances chemosensitivity of gastric cancer cells to 5-FU. *Arch. Virol.* **162,** 369–377 (2017).
66. Kadalayil, L. *et al.* Germline variation in ADAMTSL1 is associated with prognosis following breast cancer treatment in young women. *Nat. Commun.* **8,** 1632 (2017).
67. Kao, P. Y. P., Leung, K. H., Chan, L. W. C., Yip, S. P. & Yap, M. K. H. Pathway analysis of complex diseases for GWAS, extending to consider rare variants, multi-omics and interactions. *Biochim. Biophys. Acta* **1861,** 335–353 (2017).


31